

# Segmenting Images Analytically in Shape Space

Yogesh Rathi<sup>a</sup>, Samuel Dambreville<sup>b</sup>, Marc Niethammer<sup>a</sup>, James Malcolm<sup>b</sup>, James Levitt<sup>a</sup>,  
Martha E. Shenton<sup>a</sup>, Allen Tannenbaum<sup>b</sup>

<sup>a</sup>Department of Psychiatry, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA;

<sup>b</sup>School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA

## ABSTRACT

This paper presents a novel analytic technique to perform shape-driven segmentation. In our approach, shapes are represented using binary maps, and linear PCA is utilized to provide shape priors for segmentation. Intensity based probability distributions are then employed to convert a given test volume into a binary map representation, and a novel energy functional is proposed whose minimum can be analytically computed to obtain the desired segmentation in the shape space. We compare the proposed method with the log-likelihood based energy to elucidate some key differences. Our algorithm is applied to the segmentation of brain caudate nucleus and hippocampus from MRI data, which is of interest in the study of schizophrenia and Alzheimer's disease. Our validation (we compute the Hausdorff distance and the DICE coefficient between the automatic segmentation and ground-truth) shows that the proposed algorithm is very fast, requires no initialization and outperforms the log-likelihood based energy.

## 1. INTRODUCTION

Segmentation is an essential task in the analysis of medical image data as has been documented in the huge literature devoted to the subject. Of course, as in many practical problems, medical image segmentation algorithms often have to deal with poor image contrast, noise and missing boundaries. For example, consider Figure 1(a) which shows the intensity probability distribution of the caudate nucleus and its background learned *a priori* from training data. Figure 1(b) shows the likelihood of each pixel belonging to the object (caudate nucleus) based on the probabilities learned. Clearly the object and background distributions are quite close and hence insufficient to perform accurate segmentation. A typical way to deal with this problem is to incorporate prior shape knowledge in the segmentation process which brings us to the theme of this present work.

Many different methods (using both parameterized or implicit representation of shapes) have been proposed<sup>1-8</sup> to separate an object from the background using shape information. Many authors<sup>3,6,7,9,10</sup> use linear PCA to provide shape prior in their segmentation algorithms. Specifically, a few of those<sup>6,7,10</sup> use an implicit representation of shapes and do PCA on a set of signed distance functions. Recently, Dambreville *et.al*<sup>11</sup> have

---

Author information: {yogesh,marc,jlevitt,shenton@bwh.harvard.edu},  
{samuel.dambreville,malcolm,tannenba@bme.gatech.edu}

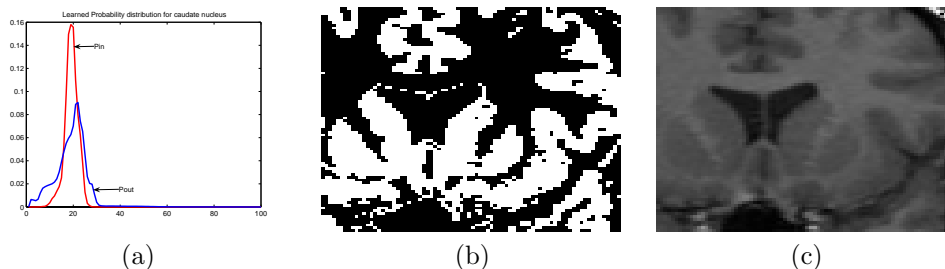


Figure 1. (a) Learned probabilities: red - caudate nucleus, blue - background (b) bright pixels are candidates for belonging to the caudate nucleus based on the probabilities learnt. (c) corresponding image slice

shown that using PCA on a set of shapes represented as binary maps provides better shape prior compared to a set of signed distance functions. In this work, we will use this implicit representation (binary map) in our segmentation algorithm, the reason for which we will elucidate below.

Most of the methods mentioned above minimize an energy functional of the form:  $E_{tot} = E_{image} + \beta E_{shape}$ , where two completely different and unrelated energy functionals ( $E_{image}$  - image energy,  $E_{shape}$  - shape energy) are linearly combined to obtain the desired segmentation. Here, one has to choose the parameter  $\beta$  which can be different for different imaging modalities, and also if one is segmenting different structures (ventricles, hippocampus etc).

The method proposed by Yezzi *et.al*<sup>7</sup> performs segmentation by separating the mean intensities (or variances) using a parametric evolution of the shape coefficients obtained by linear PCA on a set of signed distance functions. This method has the advantage of using only a single image based energy functional unlike the energy  $E_{tot}$ . It however is too restrictive in that it uses only the first two intensity moments to perform segmentation. We will use a generalized version of the image energy which will be the benchmark algorithm against which we will compare the proposed algorithm.

In this work, we perform segmentation in a novel fashion by first converting the image into a representation in shape space, and then minimizing the shape distance between the image-shape and a shape prior obtained using linear PCA. To the best of our knowledge, this is the first time segmentation has been cast as a problem of minimizing distance between two shapes using a single energy functional, whose solution can be obtained analytically and thus we regard this as the main contribution of this work. As we will try to show, this can be very useful for the segmentation of certain brain structures of interest in such diseases as schizophrenia.<sup>12</sup>

In the next section, we briefly describe the concept of linear PCA. In Section 3, we present our segmentation algorithm, and then in Section 4, we illustrate our method with some segmentation results on the left caudate nucleus and left hippocampus. We conclude the paper by giving future research directions in Section 5.

## 2. SHAPE PRIOR USING LINEAR PCA

Shape representation is an important aspect of shape learning. As mentioned earlier, we will use the binary map representation of shapes (see the paper by Dambreville *et. al.*<sup>11</sup> to get a detailed description on advantages of using a binary map representation for shapes instead of the signed distance function). A binary map representation  $\Phi$  of a shape  $S$  is obtained as follows:

$$\Phi(x) = \begin{cases} 1 & x \text{ is inside or on boundary of } S, \\ 0 & \text{else.} \end{cases} \quad (1)$$

Linear PCA (LPCA) is widely used to learn the statistical variations of a given set of data (shapes, in our case). LPCA typically assumes that all possible shapes can be written as a linear combination of a set of eigenshapes obtained by doing principal component analysis (PCA) on the training data set.<sup>6,7</sup> The eigenshapes can be obtained as follows: Let  $\beta_i$  represent the binary map corresponding to the shape  $S_i$ . The mean shape,  $\mu$ , is computed by taking the mean of the binary maps,  $\mu = \frac{1}{n} \sum \beta_i$ . The variance in shape is computed using PCA, i.e., the mean shape  $\mu$  is subtracted from each  $\beta_i$  to create a mean-offset map  $\bar{\beta}_i$ . Each such map,  $\bar{\beta}_i$ , is placed as a column vector in an  $N^d \times n$ -dimensional matrix  $M$ , where  $\beta_i \in \mathbf{R}^{N^d}$ . Using Singular Value Decomposition (SVD), the covariance matrix  $\frac{1}{n} MM^T$  is decomposed as:  $U \Sigma U^T = \frac{1}{n} MM^T$ , where  $U = [\phi_1 \phi_2 \dots \phi_n]$  is a matrix whose column vectors  $\phi_i$  represent the set of orthogonal modes of shape variation (eigenshapes) and  $\Sigma$  is a diagonal matrix of corresponding eigenvalues. An estimate of a novel shape  $\Phi$  can be obtained from  $m$  principal components using an  $m$ -dimensional vector of coefficients,  $\alpha = U_m^T (\Phi - \mu)$ , where  $U_m$  is a matrix consisting of the first  $m$  columns of  $U$ . Given the coefficients  $\alpha = [\alpha_1 \alpha_2 \dots \alpha_m]^T$ , an estimate of the shape  $\Phi$ , can be obtained as:<sup>6,7</sup>

$$\Phi = U_m \alpha + \mu = \sum_{i=1}^m \alpha_i \phi_i + \mu. \quad (2)$$

### 3. SEGMENTING IN SHAPE SPACE

For the purpose of comparison, we will use the log-likelihood based energy<sup>10</sup> for parametric evolution of the surface.<sup>7</sup> The energy is defined as follows:

$$E_{log}(\Phi_{PCA}) = - \int \log P_{in}(I(x))\Phi_{PCA}(x) + \log P_{out}(I(x))(1 - \Phi_{PCA}(x))dx, \quad (3)$$

where  $\Phi_{PCA}$  is the binary map representation of the current shape as computed from (2),  $P_{in}$  and  $P_{out}$  are probability distributions of the object and background respectively. These distributions are learned *a-priori* from the training data set using Parzen window estimation method.<sup>13</sup> Segmentation can be performed using gradient descent on the PCA coefficients  $\alpha_i$  by computing the gradient:

$$\nabla_{\alpha_i} E_{log} = \int \log \left( \frac{P_{out}(I(x))}{P_{in}(I(x))} \right) \nabla_{\alpha_i} \Phi_{PCA}(x) dx.$$

This energy will be used as a benchmark for testing the performance of the proposed method. Notice however, that the energy  $E_{log}$  is defined in the image domain.

The representation  $\Phi_{PCA}$  obtained is not a binary map, but contains values between  $[0, t]$ , where  $t \in \mathcal{R}$ . But a binary map can be easily obtained by normalizing  $\Phi_{PCA}$  to have values between 0 and 1 and then thresholding at the 0.5 level set to obtain a binary map representation. This operation does not affect the final segmentation, as can be clearly seen by re-writing  $E_{log}$  as follows:  $E_{log}(\tilde{\Phi}_{PCA}) = - \int \log P_{in}(I(x))t\tilde{\Phi}_{PCA}(x) + \log P_{out}(I(x))(1 - t\tilde{\Phi}_{PCA}(x))dx$ , with thresholding applied at  $t/2$  and  $\tilde{\Phi}_{PCA}$  is obtained by normalizing  $\Phi_{PCA}$ .

#### 3.1 The Proposed Energy in Shape Space

To perform segmentation in the shape space, an important requirement is to convert the image itself into a shape so that a meaningful distance can be computed between two shapes (one shape from the image and another from the LPCA based shape prior). A simple, yet effective technique to convert an image  $I$  into a binary shape  $\Psi$  using probability distributions is given by:

$$\Psi(x) = \begin{cases} 1 & \text{if } P_{in}(I(x)) \geq P_{out}(I(x)) , \\ 0 & \text{else .} \end{cases} \quad (4)$$

A smoother representation for  $\Psi$  can be computed using the well-known Heaviside function  $\mathbb{H}$ :<sup>7</sup>

$$\mathbb{H}(P) = \begin{cases} 1 & P > \epsilon , \\ 0 & P < -\epsilon , \\ \frac{1}{2} \{1 + \frac{P}{\epsilon} + \frac{1}{\pi} \sin(\frac{\pi P}{\epsilon})\} & \text{else,} \end{cases}$$

and  $\Psi(x) = \mathbb{H}(P_{in}(I(x)) - P_{out}(I(x)))$ . Thus,  $\Psi(x)$  is the shape map corresponding to the image  $I$  (see Figure 1(b),(c)).

Many metrics have been proposed to compute the distance between two shapes.<sup>14-16</sup> In this work, we propose to minimize the  $L^2$  distance between two shapes as follows:

$$E_{sd}(\Phi_{PCA}) = \frac{1}{2} \int (\Psi(x) - \Phi_{PCA}(x))^2 dx, \quad (5)$$

where  $\Phi_{PCA}$  is the shape prior obtained from equation (2) and  $\Psi(x)$  is computed *a-priori* as described before. Of course, one could utilize any metric of similarity between shapes proposed in the literature instead of the standard  $L^2$  metric given above in our overall framework. Our consideration for this particular choice was the convex nature of  $E_{sd}$  which has a well defined minimum, the low computational cost associated with computing  $E_{sd}$  (and its minimum) and the simplicity with which one can represent (and compute)  $\Psi(x)$ . Note that, if one were to use any parametric representation of shapes (as required for many shape metrics<sup>16</sup>), the parameterization of  $\Psi(x)$  would become extremely complex (see Figure 1(b)). Moreover, many metrics are well behaved only if the

shapes being compared are closed, connected and of genus zero.<sup>15,16</sup> However, we believe that the consideration of other possible shape metrics in order to carry out segmentation in our set-up is very important, and will be a topic for future research.

The minimum of  $E_{sd}$  will give us the desired segmentation. This can be obtained by solving the following equation to obtain each  $\alpha_i$  (the PCA coefficients):

$$\begin{aligned} 0 = \nabla_{\alpha_i} E_{sd} &= \int \left( \Psi(x) - \sum_{j=1}^m \alpha_j \phi_j - \mu \right) \phi_i(x) dx, \\ \implies \alpha_i &= \int (\Psi(x) - \mu) \phi_i(x) dx, \end{aligned} \quad (6)$$

where,  $\phi_i$  is the  $i^{th}$  eigenshape as computed in Section 2. Thus, each  $\alpha_i$  can be analytically computed and the final segmentation (shape) can be obtained from equation 2. Thus, the segmentation is just the projection of  $\Psi(x)$  in the PCA basis.

## 4. EXPERIMENTAL VALIDATION

We applied the proposed algorithm for segmenting the left caudate nucleus and the left hippocampus from MRI brain scans. The results were compared with those obtained using the log-likelihood based energy  $E_{log}$  defined earlier in (3). For both algorithms, we used the same training data set, test volumes and the top 20 PCA coefficients which contained 99 percent of the total shape variations. As a pre-processing step, the test images were registered using a nonlinear registration algorithm to a template image of the training data and segmentation was analytically obtained by projecting the corresponding shape map into the PCA basis as described earlier.

To measure the discrepancy between the segmented shape ( $S$ ) and the ground truth ( $G$ ) (obtained from hand-segmented label maps by a medical expert)\*, we use the Hausdorff distance  $H(G, S)$  that measures the maximum error between the boundary of two shapes  $G$  and  $S$ , i.e.,  $H(G, S) = \max_{g \in G} \{\min_{s \in S} \|g - s\|\}$ . We also measured the performance of our method by computing a different metric, the DICE coefficient, between  $S$  and  $G$ . The DICE coefficient is given by:

$$D(S, G) = \frac{V(S) \cap V(G)}{\frac{1}{2}(V(S) \cup V(G))},$$

where  $V$  is the volume of the shape in pixels.  $D$  measures the overlap between the automatic segmentation and ground truth.

### 4.1 Left Caudate Nucleus

For the left caudate nucleus, the training set consisted of 21 3D shapes. Segmentation was performed on 5 different test volumes not in the training set. Figure 2 shows the segmentation results for the proposed energy functional  $E_{sd}$  (5), and the log-likelihood energy  $E_{log}$  (3). Table 1 gives the Hausdorff error  $H$ , while Table 2 gives the DICE coefficient between the ground truth (GT) and the segmentation obtained using each of the methods.

**Table 1.** Hausdorff error  $H$  (mm)

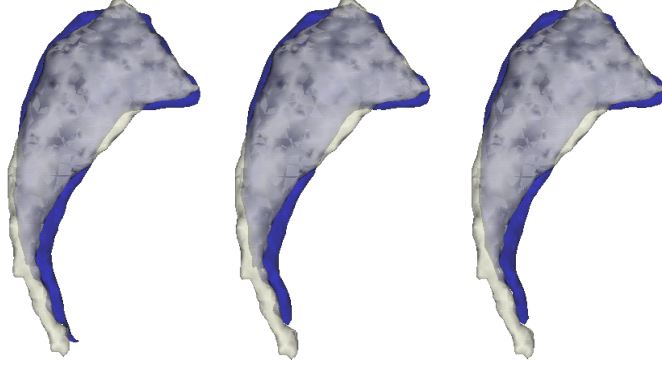
Segmentation Algorithm	N=1	N=2	N=3	N=4	N=5
Log-likelihood $E_{log}$	4.69	12.15	7.55	8.89	3.18
Shape distance $E_{sd}$	<b>2.09</b>	<b>6.15</b>	<b>4.03</b>	<b>5.56</b>	<b>2.57</b>

**Table 2.** DICE coefficient  $D$

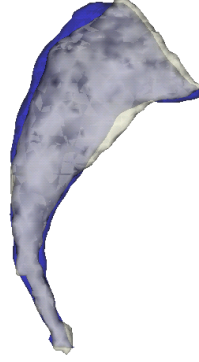
Segmentation Algorithm	N=1	N=2	N=3	N=4	N=5
Log-likelihood $E_{log}$	0.840	0.796	0.822	0.801	0.811
Shape distance $E_{sd}$	<b>0.852</b>	<b>0.825</b>	<b>0.840</b>	<b>0.828</b>	<b>0.813</b>

---

\*The training data (ground truth) was obtained from the NAMIC data repository of the Brigham and Women's Hospital, Boston, MA.



Segmentation using energy  $E_{log}$  (initial, intermediate and final surface)



Segmentation using energy  $E_{sd}$  (final surface)

Figure 2. White volume is GT. Blue volume is initialization/segmentation.

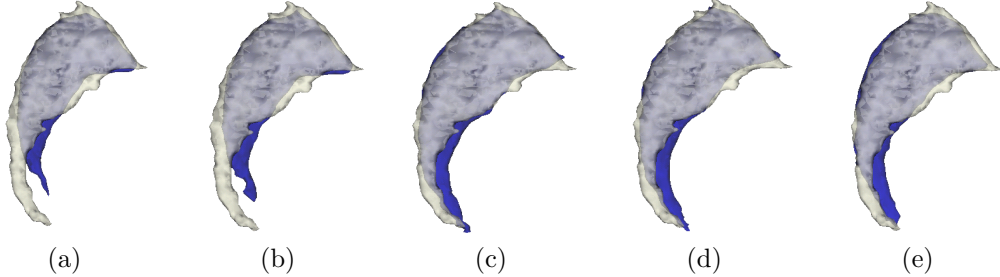


Figure 3. Caudate nucleus: White volume is GT. Blue volume is segmentation. Hausdorff error  $H$  is given in mm. (a) Far initialization, (b) Segmentation with  $E_{log}$ ,  $H = 7.55$ , (c) Close initialization, (d) Segmentation with  $E_{log}$ ,  $H = 4.14$ , (e) Segmentation with  $E_{sd}$ ,  $H = 4.03$ .

Clearly, the proposed method outperforms the log-likelihood based technique. Furthermore, the results shown here for the log-likelihood based energy were the best after many trials and many different initializations. Contrast this with the proposed method where no initialization is required. Figure 3 shows the segmentation from two different starting surfaces for the log-likelihood based technique. Clearly, the energy  $E_{log}$  gets stuck in a local minimum giving poor segmentation result, whereas the proposed shape based energy  $E_{sd}$  has a global minimum independent of the initialization. This is further illustrated in Figure 4, wherein we show the 3D isocontour lines of  $E_{sd}$  and  $E_{log}$  obtained by varying the first two PCA coefficients  $\alpha_1, \alpha_2$ . It becomes quite clear that the shape functional  $E_{sd}$  is very smooth and has a single global minimum. In contrast, the log based energy  $E_{log}$  has many local minima within which the gradient descent algorithm gets stuck leading to inaccurate segmentation.

Note that both the energy functionals (3) and (5) use exactly the same information  $(P_{in}, P_{out})$ , however, due to our formulation of the cost function in the shape space, we get much better results. Another important point to note is that, the log function is not defined when either  $P_{in}$  or being a always convex functional converges to the

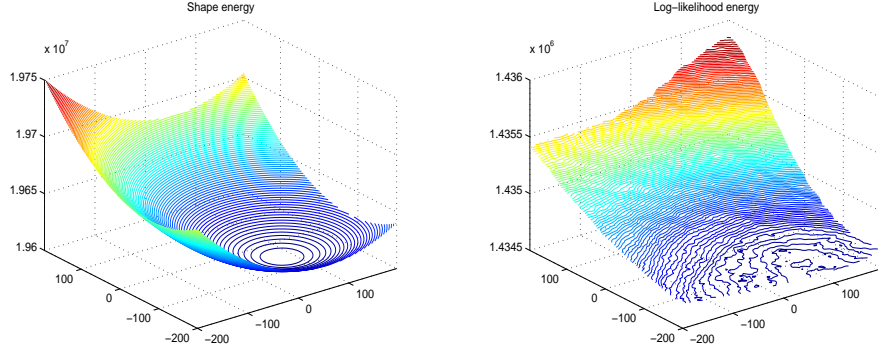


Figure 4. Isocontour plot for energy  $E_{sd}$  (left) and  $E_{log}$  (right) computed by varying the first 2 PCA coefficients within  $\pm 2$  standard deviations.

same minimum with better results  $P_{out}$  is zero, whereas no such singularity occurs in the proposed formulation (5). Furthermore, the log function descends very rapidly to  $-\infty$  as the probability goes closer to zero. Hence, the gradient contribution from points with very small probabilities dominates the total gradient (see computation of  $\nabla_{\alpha_i} E_{log}$ ). The shape energy  $E_{sd}$ , however, suffers from no such handicap.

## 4.2 Left Hippocampus

being a always convex functional converges to the same minimum with better results In this experiment, the training set contained 20 shapes, with 3 test volumes. Figure 5, Table 3 and Table 4 give the segmentation results using both methods. Once again, the shape energy performs better than the log based energy. For all the experiments (hippocampus and left caudate nucleus) using the shape energy  $E_{sd}$ , the final segmentation was obtained within 2-3 seconds on a 2.5 GHz, linux machine with 4 GB RAM. The software was written in C++ using ITK. On the other hand, for the energy  $E_{log}$  the gradient descent step size had to be tuned to ensure that the energy decreased at each iteration. Also, the number of iterations required for convergence varied from 25-50 and it took about 25-30 seconds to obtain the desired segmentation. Figure 6 gives the 3D isosurface plot of the segmented caudate nucleus and hippocampus using the 3D Slicer tool developed at the Surgical Planning lab of Brigham and Women’s hospital.

**Table 3.** Hausdorff error  $H$  (mm)

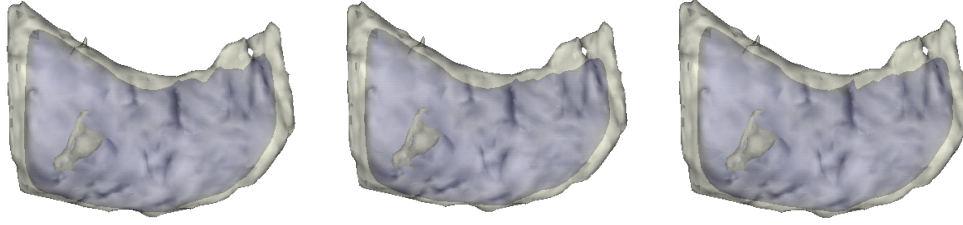
Segmentation Algorithm	N=1	N=2	N=3
Log-likelihood $E_{log}$	3.32	2.57	6.46
Shape distance $E_{sd}$	<b>2.65</b>	<b>1.76</b>	<b>3.86</b>

**Table 4.** DICE coefficient  $D$

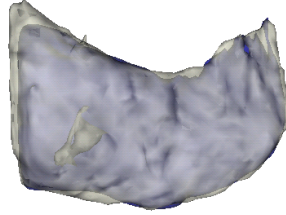
Segmentation Algorithm	N=1	N=2	N=3
Log-likelihood $E_{log}$	0.83	0.82	0.72
Shape distance $E_{sd}$	<b>0.83</b>	<b>0.83</b>	<b>0.76</b>

## 5. CONCLUSIONS AND FUTURE RESEARCH

In this work, we have presented a novel approach to perform segmentation in the shape space using  $L^2$  distance between shapes. Our preliminary results show that the proposed energy functional gives better segmentation results as compared to the more standard log-based energy. The segmentation is obtained completely automatically with no need for initialization in contrast to many techniques reported in the literature. Also, this is the first time an analytic technique has been proposed to obtain segmentation using a novel energy formulation. Nevertheless, the proposed method has a few drawbacks. The segmentation results are sensitive to the alignment of the test images to the template image. Error in the registration process can affect the quality of segmentation obtained. In case the registration of images is not accurate enough, the proposed method can at the very least serve as a very quick initialization process for a more sophisticated algorithm. Note that, no user interaction whatsoever is required to obtain the segmentation, which is a step forward towards completely automatic segmentation algorithms.



Segmentation using  $E_{log}$  (initial, intermediate and final)



Segmentation using  $E_{sd}$  (final)

Figure 5. White volume is GT. Blue volume is initialization/segmentation.

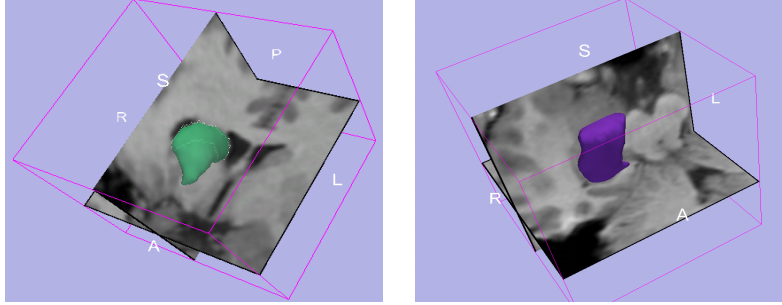


Figure 6. 3D view of the segmented isosurface of the caudate nucleus (left) and hippocampus (right) using shape distance  $E_{sd}$ .

Future research includes exploring other metrics for computing shape distance as well as the possible use of different representations of shape along with using other shape learning techniques like kernel PCA. Moreover, we plan to test the methodology on more data sets as well as different structures. For example, this segmentation could be very useful for the analysis of the prostate in surgical planning. Presumably the method could also be included in the pipeline for the study of the statistics of populations in some of the studies being conducted in trying to utilize shape information of various brain structures to differentiate normal from schizophrenic patients. In summary, we believe that we have proposed an efficient and robust segmentation approach which hopefully will find its ways into a number of medical applications.

## 6. ACKNOWLEDGMENTS

The authors would like to thank Dr. Sylvain Bouix of Psychiatry Neuroimaging lab, Harvard Medical School, for fruitful discussions on this topic.

## REFERENCES

1. Nain, D., Haker, S., Bobick, A., Tannenbaum, A.: Shape-driven 3d segmentation using spherical wavelets. In: MICCAI. (2006)
2. Xu, C., Pham, D., Prince, J.L.: Medical image segmentation using deformable models. In: SPIE. (2000) 129–174
3. Cootes, T., Taylor, C., Cooper, D., Graham, J.: Active shape models, their training and application. In: Computer Vision and Image Understanding. Volume 61. (1995) 38–59

4. Faugeras, O., Gomes, J.: Dynamic shapes of arbitrary dimension: the vector distance functions. In Cipolla, R., Martin, R., eds.: Proc. of 9th IMA Conference on Mathematics of Surfaces. The Mathematics of Surfaces IX, Springer (2000)
5. Cremers, D., Kohlberger, T., Schnorr, C.: Nonlinear shape statistics in mumford-shah based segmentation. In: 7th ECCV '02. Volume 2351. (2002) 93–108
6. Leventon, M., Grimson, E., Faugeras, O.: Statistical shape influence in geodesic active contours. In: Proc. CVPR, IEEE (2000) 1316–1324
7. Tsai, A., Yezzi, A., Wells, W., Tempany, C., Tucker, D., Fan, A., Grimson, E., Willsky, A.: A shape-based approach to the segmentation of medical imagery using level sets. IEEE Trans. on Medical Imaging **22** (2003) 137–153
8. Dambreville, S., Rath, Y., Tannenbaum, A.: Shape-based approach to robust image segmentation using kernel PCA. In: CVPR. (2006)
9. Wang, Y., Staib, L.: Boundary finding with correspondence using statistical shape models. In: Proc. CVPR. (1998) 338–345
10. Rousson, M., Cremers, D.: Efficient Kernel Density Estimation of Shape and Intensity Priors for Level Set Segmentation. In: Medical Image Computing and Computer Assisted Intervention. Volume 1. (2005) 757–764
11. Dambreville, S., Rath, Y., Tannenbaum, A.: A shape-based approach to robust image segmentation. In: Intl. Conf. Image Analysis and Recognition. (2006)
12. S.Vetsa, Styner, M., Pizer, S., Lieberman, J., Gerig, G.: Caudate shape discrimination in schizophrenia using template-free non-parametric tests. In: MICCAI. (2003) 661–669
13. Duda, R., Hart, P., Stork, D. In: Pattern Classification. Wiley-Interscience (2001)
14. Yezzi, A., Mennucci, A.: Metrics in the space of curves. <http://arxiv.org/abs/math.DG/0412454> (Accessed in July 2005) (2004)
15. Michor, P., Mumford, D.: Riemannian geometries of space of plane curves. (<http://front.math.ucdavis.edu/math.DG/0312384>. (Accessed in July 2005))
16. Srivastava, A., Joshi, S., Mio, W., Liu, X.: Statistical shape analysis: Clustering, learning and testing. Trans. PAMI (2005)